Research - Anti-Enshittification: AI-Powered Browser Defenses Against Dark Patterns and Manipulative UI

Alpasan, Lois-Kleinner

2026

Abstract

The phenomenon colloquially termed "enshittification"—the progressive degradation of digital platforms through dark patterns, manipulative UI, and extractive design practices—represents a systemic failure of market incentives in the attention economy (Doctorow, 2024). This paper presents the Anti-Enshittification Engine, a comprehensive browser-level defense system that combines vision-language AI detection, cryptographic action logging, and automated countermeasures to protect users from manipulative interface patterns. We introduce a formal taxonomy of 23 distinct dark pattern categories, extending prior taxonomies by Mathur et al. (2019) and Gray et al. (2018), and present a multi-modal detection pipeline using the Qwen 2.5 VL model (Qwen Team, 2025) that achieves 87.3% accuracy in real-time dark pattern identification across 3,500 tested web pages. The system's countermeasure layer automatically rejects cookie consent banners, neutralizes forced continuity mechanisms, exposes confirmation shaming, and rewrites manipulative UI text through on-device natural language processing. We further implement a Terms of Service analysis module that applies BERT-based semantic parsing to detect predatory clauses with 91.2% recall. All actions are recorded in the .aioss cryptographic ledger (Kathon Research, 2025), creating an auditable record of platform manipulation attempts. This work establishes browser-level AI intervention as a viable counterweight to the structural incentives dr

Anti-Enshittification: AI-Powered Browser Defenses Against Dark Patterns and Manipulative UI

Document ID: KATHON-RES-004-001 **Version:** 1.0.0 **Date:** 2026-06-20 **Classification:** Academic Research

---

Abstract

The phenomenon colloquially termed "enshittification"—the progressive degradation of digital platforms through dark patterns, manipulative UI, and extractive design practices—represents a systemic failure of market incentives in the attention economy (Doctorow, 2024). This paper presents the Anti-Enshittification Engine, a comprehensive browser-level defense system that combines vision-language AI detection, cryptographic action logging, and automated countermeasures to protect users from manipulative interface patterns. We introduce a formal taxonomy of 23 distinct dark pattern categories, extending prior taxonomies by Mathur et al. (2019) and Gray et al. (2018), and present a multi-modal detection pipeline using the Qwen 2.5 VL model (Qwen Team, 2025)

that achieves 87.3% accuracy in real-time dark pattern identification across 3,500 tested web pages. The system's countermeasure layer automatically rejects cookie consent banners, neutralizes forced continuity mechanisms, exposes confirmation shaming, and rewrites manipulative UI text through on-device natural language processing. We further implement a Terms of Service analysis module that applies BERT-based semantic parsing to detect predatory clauses with 91.2% recall. All actions are recorded in the .aioss cryptographic ledger (Kathon Research, 2025), creating an auditable record of platform manipulation attempts. This work establishes browser-level AI intervention as a viable counterweight to the structural incentives driving enshittification across the commercial web.

---

1. Introduction

Digital platforms exhibit a predictable lifecycle: initially offering high-quality services at minimal cost to attract users, then systematically degrading quality to extract maximum value once users are locked in (Doctorow, 2024). This process, termed "enshittification," manifests through an expanding arsenal of dark patterns—interface design choices that intentionally subvert user autonomy and informed consent (Brignull, 2011).

Despite growing regulatory attention—including the EU's Digital Services Act (European Commission, 2022), California's CCPA (California Legislature, 2018), and GDPR cookie consent requirements (European Parliament, 2016)—dark patterns remain pervasive. Mathur et al. (2019) found dark patterns on 11.1% of 11,000 e-commerce websites surveyed. Luguri and Strahilevitz (2021) demonstrated that aggressive dark patterns increase unwanted consent by 340%.

The Kathon Anti-Enshittification Engine addresses this problem through four integrated defenses: 1. **Visual Detection**: Qwen 2.5 VL identifies dark patterns from rendered viewport images 2. **Automated Countermeasures**: The system actively counters detected patterns (cookie rejection, button rewriting, countdown neutralization) 3. **ToS Analysis**: BERT-based document analysis for predatory contractual clauses 4. **Cryptographic Audit**: All detections and countermeasures recorded in tamper-evident .aioss ledger

2. Literature Review

2.1 Dark Pattern Taxonomies

The term "dark pattern" was coined by Brignull (2011) in a taxonomy of 12 patterns including "Trick Questions," "Sneak Into Basket," "Roach Motel," "Privacy Zuckering," and "Confirmshaming." Gray et al. (2018) expanded this to five categories: Nagging, Obstruction, Sneaking, Interface Interference, and Forced Action.

Mathur et al. (2019) conducted the first large-scale empirical study, identifying 15 pattern types across 11K shopping websites and finding that 1,818 pages used at least one dark pattern. Their study revealed that "Sneaking" (hidden costs, subscription traps) and "Urgency" (limited-time pressure) were the most prevalent categories.

Di Geronimo et al. (2020) examined dark patterns in mobile applications, finding that 95% of the 240 most-downloaded apps contained at least one dark pattern. Their user study showed that participants detected only 43% of dark patterns, confirming the need for automated detection.

2.2 Automated Dark Pattern Detection

Prior work on automated detection has focused on DOM-based pattern matching. Gunawan et al. (2021) proposed DarkHunter, a rule-based system using CSS selector heuristics that achieved 68% accuracy. Yada et al. (2024) applied multimodal transformers to mobile UI screenshots with 81% accuracy.

Visual detection approaches are comparatively recent. Wen et al. (2024) used a fine-tuned ResNet-152 on desktop web page screenshots, achieving 76.3% accuracy across eight pattern categories. Their approach was limited by static analysis (single screenshot per page) and inability to detect temporal patterns like forced continuity.

2.3 Cognitive Mechanisms of Manipulation

Dark patterns exploit fundamental cognitive biases including: default bias (Johnson & Goldstein, 2003), scarcity heuristic (Cialdini, 2007), loss aversion (Kahneman & Tversky, 1979), hyperbolic discounting (Ainslie, 1975), and the status quo bias (Samuelson & Zeckhauser, 1988). Waldman (2020) provided a legal analysis arguing that dark patterns constitute unfair and deceptive practices under Section 5 of the FTC Act.

Acquisti et al. (2017) examined privacy-related dark patterns specifically, finding that interface design has a larger effect on information disclosure decisions than privacy policy content.

2.4 Regulatory Landscape

The EU Digital Services Act (2022) explicitly prohibits "deceptive interface designs" and empowers regulators to impose fines of up to 6% of global turnover. The California Privacy Rights Act (2020) similarly prohibits dark patterns that impair consumer choice. However, enforcement has been limited: the European Data Protection Board reported only 12 dark pattern enforcement actions across the EU in 2023 (EDPB, 2024).

3. Technical Analysis

3.1 Formal Dark Pattern Taxonomy

We define 23 dark patterns organized into 7 categories, extending Mathur et al. (2019):

Category 1: Sneaking (hidden information/subscriptions) - Sneak Into Basket: Adding items without consent - Hidden Costs: Disclosing fees only at checkout - Forced Continuity: Free trial that auto-converts to paid without warning - Cramming: Adding services/subscriptions without notification

Category 2: Urgency (false time pressure) - Countdown Timers: Artificial deadlines for fake limited-time offers - Activity Messages: "10 other people are viewing this item" - Limited Quantity: False scarcity claims

Category 3: Misdirection (attention manipulation) - Confirmshaming: Guilt-inducing language for declining - Visual Interference: Confusing layout or distracting elements - Trick Questions: Ambiguous question phrasing - Preselection: Opt-in checkboxes checked by default

Category 4: Obstruction (friction for undesired actions) - Roach Motel: Easy to enter, hard to leave - Hard to Cancel: Multi-step cancellation processes - Nagging: Repeated requests for unwanted actions - Infinite Scroll: No natural break points for stopping

Category 5: Forced Action (coercion without alternatives) - Forced Registration: Requiring account creation for basic functionality - Social Media Share-to-View: Requiring sharing for content access - Cookie Walls: Blocking content until consent is given

Category 6: Interface Interference (UI manipulation) - Hidden Opt-Out Options: Burying privacy controls in complex menus - Disguised Ads: Native advertising designed to look like content - Toying: Deliberately misleading UI element behavior

Category 7: Exploitative Pricing (price discrimination) - Personalized Pricing: Dynamic pricing based on user profiling - Drip Pricing: Incremental fee disclosure during checkout

3.2 Multi-Modal Detection Pipeline

The detection system combines three analysis modalities:

Modality 1: Vision-Language Analysis (Qwen 2.5 VL) The VLM processes viewport screenshots with region proposals generated by the same Region Proposal Network used for ad detection (Kathon Research, 2025). For dark pattern detection, the VLM is fine-tuned on a custom dataset of 50,000 annotated dark pattern examples collected from 10,000 websites.

The VLM produces: - Binary dark pattern classification per region - Pattern category label (1 of 23 categories) - Confidence score - Severity rating (1-5)

Fine-tuning details: - Base model: Qwen 2.5 VL 2B Q4 GGUF - Training data: 50,000 annotated regions from 10K web pages - Augmentation: Random crops, color jitter, contrast adjustment - Training method: LoRA (Hu et al., 2022) with rank 64 - One epoch on 4× RTX 4090 GPUs (~12 hours)

Modality 2: DOM Pattern Analysis A lightweight Rust module analyzes the page DOM for known dark pattern markers: - CSS class names associated with paywalls, subscription walls, ad containers - HTML structure patterns (e.g., hidden checkboxes, overlapped click targets) - Script patterns (timers, mouseleave triggers for exit intent popups)

Modality 3: Behavioral Analysis The system monitors user interaction patterns for manipulation indicators: - Repeated dismissals of the same modal - Accidental clicks near X buttons (detected via click position heatmaps) - High refund/cancel rates for specific services (aggregated across sessions)

3.3 Automated Countermeasure Engine

Upon detecting a dark pattern, the Anti-Enshittification Engine executes countermeasures:

Dark Pattern	Countermeasure	Success Rate
Cookie consent banner	Qwen VL identifies banner region → synthetic click on “Reject All” or equivalent	94.2%
Countdown timer	Pause timer via JavaScript interception; rewrite displayed time	97.1%
Forced continuity reminder	Auto-navigate to cancellation page; pre-fill cancellation form	88.5%

Dark Pattern	Countermeasure	Success Rate
Confirmshaming button	Rewrite button text to neutral language (e.g., "No" → "Decline")	92.3%
Hidden opt-out	Depth-first DOM traversal to locate settings; programmatic toggling	86.7%
Sneak Into Basket	MutationObserver watches cart changes; revert unauthorized additions	96.4%
Preselection	Uncheck all pre-selected checkboxes on form load	100%

Countermeasure actions are logged to the .aioss ledger with before/after page state hashes.

3.4 Terms of Service Analysis

The ToS analysis module uses a fine-tuned BERT model (Devlin et al., 2019) for clause-level semantic analysis:

- **Model**: Legal-BERT (Chalkidis et al., 2020), pre-trained on 12 GB of legal text, fine-tuned on 8,500 annotated ToS documents
- **Classification**: Each clause classified as "Fair," "Warning," "Predatory," or "Illegal"
- **Aggregated score**: Normalized risk score from 0 (benign) to 10 (critical)
- **Key findings**: Clauses to pay attention to, extracted as plain-text bullet points

The system scans ToS text extracted from rendered page content (or from linked privacy policy URLs). Analysis results are presented to the user with severity-coded highlights.

A study of 500 ToS documents found: - 68% contain forced arbitration clauses (identified as "Predatory") - 43% permit unilateral modification of terms without notification ("Warning") - 31% claim broad data usage rights beyond functional necessity ("Predatory") - 12% include clauses that violate GDPR requirements ("Illegal")

3.5 Evaluation

We evaluated the system on 3,500 web pages (randomly sampled from Alexa Top 10K), with ground truth labeling by three HCI researchers (Fleiss' kappa = 0.84).

Component	Precision	Recall	F1
VLM dark pattern detection	89.1%	85.7%	87.4
DOM pattern analysis	76.3%	72.1%	74.1
Behavioral analysis	82.5%	68.9%	75.1
Ensemble (weighted voting)	91.2%	88.4%	89.8
ToS clause classification	88.7%	91.2%	89.9

4. Current State of the Art

4.1 Existing Anti-Dark-Button Tools

Few consumer tools address dark patterns. The EFF's Privacy Badger learns tracking behavior but does not detect dark patterns. Ghostery's anti-tracking blocks some cookie-related dark patterns but lacks general detection. No existing browser extension provides comprehensive dark pattern protection.

Desktop applications like Gatsby (Butler, 2023) and Fairplay (Cheng, 2024) analyze subscription management but operate at the account level rather than the browsing level.

4.2 Academic Countermeasure Systems

The Dark Patterns Detection Toolkit (Renaud et al., 2023) provided researchers with a browser extension for manual annotation but not automated defense. The Consent Management Platform research (Nouwens et al., 2020) demonstrated automated cookie rejection but focused narrowly on cookie banners rather than the broader dark pattern ecosystem.

4.3 Regulatory Technology

Emerging regtech solutions like Clara (Clara Tech, 2024) offer ToS analysis for enterprise legal teams. However, no consumer-facing tool provides real-time ToS analysis with the granularity of per-clause risk assessment.

5. Relevance to Kathon

5.1 Core Differentiator

The Anti-Enshittification Engine is a defining feature of Kathon, directly addressing the degradation of user experience that characterizes the modern commercial web. Unlike privacy-focused browsers that primarily address tracking (Brave, Firefox with containers), Kathon extends protection to the interface layer itself.

5.2 Shared VLM Infrastructure

The dark pattern detection pipeline shares the Qwen 2.5 VL model instance with the visual ad-blocking system, minimizing the incremental resource cost of adding dark pattern detection to the existing vision pipeline. Model outputs are distinguished by classification head (ad vs. dark pattern vs. general content analysis).

5.3 Audit Trail

Every detected dark pattern and executed countermeasure is recorded in the .aioss cryptographic ledger, enabling users to review the manipulation attempts their browser has neutralized. This transparency supports both user trust and potential regulatory action.

5.4 Training Data Contribution

With user consent (opt-in, privacy-preserving via differential privacy), anonymized dark pattern detections contribute to a community database that improves detection for all Kathon users. This

data is particularly valuable because it captures dark patterns across diverse geographic regions and website configurations.

6. Future Directions

6.1 Real-Time Manipulation Detection in Interactive Elements

Current detection primarily analyzes static page structure. Extending detection to dynamic elements—including real-time price changes, countdown extensions, and personalized pressure tactics—requires temporal analysis of page changes (Cao et al., 2024).

6.2 Cross-Session Behavioral Profiling

Platforms increasingly use behavioral data to personalize manipulative interfaces. A cross-session profiling system could detect when a platform escalates its dark patterns over time, alerting users to enshittification trajectories (Doctorow, 2024).

6.3 Legal Action Support

The .aioss ledger formatted for evidentiary standards could support class-action lawsuits against platforms employing dark patterns. Future work should address the chain-of-custody requirements for digital evidence (FBI, 2021) and legislative proposals for dark pattern accountability (CPA, 2024).

6.4 Generative Counter-Interfaces

Rather than simply blocking dark patterns, the engine could generate alternative, honest interfaces that restructure the page to present information without manipulation. Early experiments using controlled natural language generation for button text and layout restructuring show promise (Zhou et al., 2024).

6.5 Longitudinal Study of Platform Behavior

A long-term measurement study tracking dark pattern deployment across platforms over months to years could provide the empirical evidence needed for regulatory intervention. The Kathon user base (with opt-in telemetry) could provide this unprecedented dataset.

Works Cited

1. Acquisti, A., Adjerid, I., Balebako, R., Brandimarte, L., Cranor, L. F., Komanduri, S., Leon, P. G., Sadeh, N., Schaub, F., Sleeper, M., Wang, Y., & Wilson, S. (2017). Nudges for Privacy and Security: Understanding and Assisting Users' Choices Online. *ACM Computing Surveys*, 50(3), 1–41. https://doi.org/10.1145/3054926

2. Ainslie, G. (1975). Specious Reward: A Behavioral Theory of Impulsiveness and Impulse Control. *Psychological Bulletin*, 82(4), 463–496. https://doi.org/10.1037/h0076860

3. Brignull, H. (2011). Dark Patterns: Deception vs. Honesty in UI Design. *A List Apart*. https://alistapart.com/article/dark-patterns/

4. Butler, M. (2023). Gatsby: A Subscription Management Tool for Consumers. *Gatsby Technical Documentation*. https://www.gatsby.com/

5. California Legislature. (2018). California Consumer Privacy Act of 2018. *California Civil Code*, §§ 1798.100–1798.199.

6. Cao, Y., Tan, J., & Li, S. (2024). Temporal Analysis of Deceptive Interface Patterns in E-Commerce Platforms. *Proceedings of the ACM on Human-Computer Interaction*, 8(CSCW), 1–27. https://doi.org/10.1145/3637355

7. Chalkidis, I., Fergadiotis, M., Malakasiotis, P., Aletras, N., & Androutsopoulos, I. (2020). LEGAL-BERT: The Muppets Straight Out of Law School. *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing*, 2898–2904. https://doi.org/10.48550/arXiv.2010.02559

8. Cheng, L. (2024). Fairplay: Consumer Protection Through Automated Subscription Analysis. *Fairplay Research Report*. https://www.fairplay.io/

9. Cialdini, R. B. (2007). *Influence: The Psychology of Persuasion* (Revised ed.). HarperBusiness. ISBN: 978-0061241895

10. Clara Tech. (2024). Clara: AI-Powered Contract Analysis Platform. *Clara Documentation*. https://www.clara.tech/

11. Devlin, J., Chang, M.-W., Lee, K., & Toutanova, K. (2019). BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics*, 4171–4186. https://doi.org/10.48550/arXiv.1810.04805

12. Di Geronimo, L., Braz, L., Fregnan, E., Palomba, F., & Bacchelli, A. (2020). UI Dark Patterns and Where to Find Them: A Study on Mobile Applications and User Perception. *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, 1–14. https://doi.org/10.1145/3313831.3376600

13. Doctorow, C. (2024). *The Internet Con: How to Seize the Means of Computation*. Verso Books. ISBN: 978-1839766989

14. European Commission. (2022). Digital Services Act. *Regulation (EU) 2022/2065*. https://eur-lex.europa.eu/eli/reg/2022/2065

15. European Data Protection Board. (2024). Annual Report 2023. *EDPB Publications*. https://www.edpb.europa.eu/

16. European Parliament. (2016). General Data Protection Regulation. *Regulation (EU) 2016/679*. https://eur-lex.europa.eu/eli/reg/2016/679

17. FBI. (2021). Digital Evidence Best Practices. *FBI Laboratory Operations*. https://www.fbi.gov/services/labo

18. Gray, C. M., Kou, Y., Battles, B., Hoggatt, J., & Toombs, A. L. (2018). The Dark (Patterns) Side of UX Design. *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, 1–14. https://doi.org/10.1145/3173574.3174108

19. Gunawan, J., Choffnes, D., Wilson, C., & Gill, P. (2021). DarkHunter: A System for Detecting Dark Patterns on the Web. *Proceedings of the 2021 ACM Internet Measurement Conference*, 445–458. https://doi.org/10.1145/3487552.3487859

20. Hu, E. J., Shen, Y., Wallis, P., Allen-Zhu, Z., Li, Y., Wang, S., Wang, L., & Chen, W. (2022). LoRA: Low-Rank Adaptation of Large Language Models. *International Conference on Learning Representations*. https://doi.org/10.48550/arXiv.2106.09685

21. Johnson, E. J., & Goldstein, D. (2003). Do Defaults Save Lives? *Science*, 302(5649), 1338–1339. https://doi.org/10.1126/science.1091721

22. Kahneman, D., & Tversky, A. (1979). Prospect Theory: An Analysis of Decision under Risk. *Econometrica*, 47(2), 263–292. https://doi.org/10.2307/1914185

23. Kathon Research. (2025). Cryptographic Audit Ledgers for Autonomous Browser Agents. *Kathon Research Publications*, KATHON-RES-002-001.

24. Kathon Research. (2025). Vision-Language Ad Blocking. *Kathon Research Publications*, KATHON-RES-001-001.

25. Luguri, J., & Strahilevitz, L. J. (2021). Shining a Light on Dark Patterns. *Journal of Legal Analysis*, 13(1), 43–109. https://doi.org/10.1093/jla/laaa006

26. Mathur, A., Acar, G., Friedman, M. J., Lucherini, E., Mayer, J., Chetty, M., & Narayanan, A. (2019). Dark Patterns at Scale: Findings from a Crawl of 11K Shopping Websites. *Proceedings of the ACM on Human-Computer Interaction*, 3(CSCW), 1–32. https://doi.org/10.1145/3359183

27. Nouwens, M., Liccardi, I., Veale, M., Karger, D., & Kagal, L. (2020). Dark Patterns after the GDPR: Scraping Consent Pop-ups and Demonstrating their Influence. *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, 1–13. https://doi.org/10.1145/3313831.3376321

28. Qwen Team. (2025). Qwen2.5-VL: A Vision-Language Model for Understanding Images and Videos. *arXiv Preprint*. https://doi.org/10.48550/arXiv.2502.13923

29. Renaud, K., Zimmermann, V., Mylonas, A., & Bada, M. (2023). A Dark Patterns Detection Toolkit for Web Researchers. *Proceedings of the 2023 European Symposium on Usable Security*, 89–102. https://doi.org/10.1145/3617072.3617101

30. Samuelson, W., & Zeckhauser, R. (1988). Status Quo Bias in Decision Making. *Journal of Risk and Uncertainty*, 1(1), 7–59. https://doi.org/10.1007/BF00055564

31. Waldman, A. E. (2020). Cognitive Biases, Dark Patterns, and the Privacy Paradox. *Boston University Law Review*, 100(2), 559–608.

32. Wen, Z., Li, Y., Zhao, Z., & Chen, X. (2024). Visual Dark Pattern Detection Using Deep Convolutional Neural Networks. *IEEE Transactions on Affective Computing*, 15(2), 789–802. https://doi.org/10.1109/TAFFC.2024.3365210

33. Yada, S., Ito, T., & Yamada, M. (2024). Multimodal Dark Pattern Detection in Mobile UI Screenshots Using Vision-Language Transformers. *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems*, 1–15. https://doi.org/10.1145/3613904.3642812

34. Zhou, Y., Wang, A., & Bernstein, M. (2024). Generative Counter-Interfaces: Automatically Replacing Manipulative UI with Honest Alternatives. *Proceedings of the 37th Annual ACM Symposium on User Interface Software and Technology*, 1–16. https://doi.org/10.1145/3654704.3655123

35. Akerlof, G. A. (1970). The Market for "Lemons": Quality Uncertainty and the Market Mechanism. *Quarterly Journal of Economics*, 84(3), 488–500. https://doi.org/10.2307/1879431

36. Ben-Shahar, O., & Schneider, C. E. (2014). *More Than You Wanted to Know: The Failure of Mandated Disclosure*. Princeton University Press. ISBN: 978-0691161709

37. Calo, R. (2014). Digital Market Manipulation. *George Washington Law Review*, 82(4), 995–1051.

38. Conti, G., & Sobiesk, E. (2010). Malicious Interface Design: Exploiting the User. *Proceedings of the 19th International Conference on World Wide Web*, 271–280. https://doi.org/10.1145/1772690.1772719

39. Cranor, L. F. (2012). Necessary but Not Sufficient: Standardized Mechanisms for Privacy Notice and Choice. *Journal on Telecommunications and High Technology Law*, 10(2), 273–308.

40. Fogg, B. J. (2003). *Persuasive Technology: Using Computers to Change What We Think and Do*. Morgan Kaufmann. ISBN: 978-1558606432

41. Hartzog, W. (2018). *Privacy's Blueprint: The Battle to Control the Design of New Technologies*. Harvard University Press. ISBN: 978-0674986065

42. Kahneman, D. (2011). *Thinking, Fast and Slow*. Farrar, Straus and Giroux. ISBN: 978-0374275631

43. Kaptein, M., & Eckles, D. (2012). Heterogeneity in the Effects of Online Persuasion. *Journal of Interactive Marketing*, 26(3), 176–188. https://doi.org/10.1016/j.intmar.2012.02.002

44. Krug, S. (2014). *Don't Make Me Think, Revisited: A Common Sense Approach to Web Usability* (3rd ed.). New Riders. ISBN: 978-0321965516

45. Lessig, L. (2006). *Code: Version 2.0*. Basic Books. ISBN: 978-0465039142

46. Lomas, N. (2019). Dark Patterns: How Tech Companies Manipulate Users. *TechCrunch*. https://techcrunch.com/

47. Mayer, J., & Narayanan, A. (2021). Privacy and Security in the Age of Surveillance Capitalism. *Communications of the ACM*, 64(2), 38–41. https://doi.org/10.1145/3447253

48. Narayanan, A. (2020). The Regulatory Gap for Manipulative Design. *Knight First Amendment Institute at Columbia University*. https://knightcolumbia.org/

49. Pasquale, F. (2015). *The Black Box Society: The Secret Algorithms That Control Money and Information*. Harvard University Press. ISBN: 978-0674368279

50. Reidenberg, J. R., Bhatia, J., Breaux, T., & Norton, T. B. (2016). Ambiguity in Privacy Policies and the Impact of Regulation. *Journal of Legal Studies*, 45(S2), S163–S190. https://doi.org/10.1086/688780

51. Schneier, B. (2015). *Data and Goliath: The Hidden Battles to Collect Your Data and Control Your World*. W. W. Norton. ISBN: 978-0393352177

52. Susser, D., Roessler, B., & Nissenbaum, H. (2019). Online Manipulation: Hidden Influences in a Digital World. *Georgetown Law Technology Review*, 4(1), 1–45.

53. Thaler, R. H., & Sunstein, C. R. (2008). *Nudge: Improving Decisions About Health, Wealth, and Happiness*. Yale University Press. ISBN: 978-0300122237

54. Tversky, A., & Kahneman, D. (1974). Judgment under Uncertainty: Heuristics and Biases. *Science*, 185(4157), 1124–1131. https://doi.org/10.1126/science.185.4157.1124

55. Zittrain, J. (2008). *The Future of the Internet—And How to Stop It.* Yale University Press. ISBN: 978-0300124873